Choosing the Best Server for a Data Center: The Importance of Workload Weighting

Sulav Malla, Ken Christensen
Department of Computer Science and Engineering
University of South Florida, Tampa, USA
sulavmalla@mail.usf.edu, christen@cse.usf.edu

***Abstract*—Power consumption and throughput as a function of utilization are key measures for understanding the performance-energy trade-offs of a server. Existing energy efficiency metrics use these measures in various ways. In this paper, we demonstrate the importance of taking workload characteristics into account such that the best server can be selected on a case by case basis. The best server will meet the performance demands of the workload with the lowest possible energy consumption. We rigorously evaluate existing metrics, with and without workload weighting, using publicly available SPECpower benchmark data for actual servers and a publicly available Google cluster usage data. We show that using unweighted metrics to select a server can result in selecting the wrong server ensuing in higher operational energy use, the difference being as high as 20% additional energy use. Weighting power consumption of a server by the workload distribution results in a metric that is linear and reliable in ranking of servers. An analysis of server rankings for two different existing metrics shows that our weighted metric is statistically significantly better than existing metrics. The results in this paper can help data center operators to quantify the importance of workload weighting to help them select the best server to reduce operational energy costs.**

***Index Terms*—Data center, energy efficiency, energy proportionality, metrics, workload.**

I. Introduction

Greening the data center is an active area of research. Data center electricity use in the U.S. was 70 billion kWh in 2014 with a projection to reach 73 billion kWh by 2020 [14], a significant (about 2%) portion of the entire U.S. electricity consumption. Given the massive operational cost associated with data center electricity use, it is in the interest of data center operators, as well as the society as a whole, to reduce energy use. Work has been done to make servers used in data centers more energy efficient. We require metrics in order to compare energy efficiency between servers, observe progress over time, and/or predict energy consumption of servers. Most of the server energy efficiency metrics only consider peak performance and peak power consumption which typically happens when a server is fully utilized. However, various studies of actual server deployments in different data centers [3], [14] have found that servers are rarely 100% utilized with typical average server utilization being in the range of 10% to 50% depending on the type and deployment size of the data center. Hence, an energy efficiency metric should consider the power and performance over the entire utilization range. The SPECpower_ssj2008 [16] benchmark (henceforth referred to as SPECpower) is a step in this direction.

Standard Performance Evaluation Corporation (SPEC) came up with the SPECpower benchmark after a need for a standard server energy efficiency metric was felt. The benchmark loads a server with a server-side-Java (ssj) application to exercise CPU, cache, memory, and disk to find the peak throughput. The server is then loaded from 0% to 100% of the peak throughput in steps of 10% and corresponding power is measured at these 11 different utilization levels. The benchmark then calculates the overall rating for a server as the ratio of sum of performance and sum of power consumption.

It is well understood that the benchmarking environment for a server needs to match the actual operational environment for the results to be meaningful, or reliable. The SPEC Power and Performance Methodology [15] states, "Unless the workload and configurations of the benchmarked solution match your planned solution, it could be very misleading to assume that a benchmark result will equate to reality in a production data center." While hardware vendors report the SPEC overall rating, a data center operator must take their workload characteristics into account. For example, if servers in a data center are more likely to be at 20% utilization and very rarely at 100% utilization, energy efficiency at 20% utilization should matter more. Thus, we need a metric that takes the workload characteristics of a data center into account. A key question that has not been answered is, what is the impact on expected energy savings if workload characteristics are not considered? This paper answers this important question. Key contributions from our paper are:

- A rigorous empirical study of the reliability of key energy efficiency metrics using SPECpower benchmark data for 42 different servers given Google cluster usage workload.
- A demonstration of the importance of weighting existing energy efficiency metric with workload to find the server with the best power-performance for a given workload.
- A new metric that uses workload weighting to achieve a near perfect energy efficiency ranking of servers for a given workload.

II. Review of Existing Server Metrics

Power consumption of a server as a function of utilization characterizes the power use of a server, while throughput of the server as a function of utilization characterizes its performance. Throughout this paper, we use $0 \leq u \leq 1$ to denote utilization of a server, $P(u)$ to represent the power-utilization

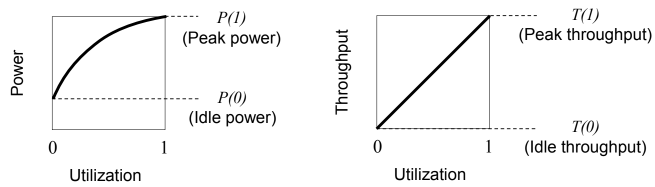

Fig. 1. Power consumption and throughput as a function of server utilization.

function, $T(u)$ for the throughput-utilization function, and the ratio, $P(u)/T(u)$, to be the energy efficiency of a server at utilization u. Figure 1 shows the $P(u)$ and $T(u)$ curves for an example server. Actual servers typically have peak power or peak performance at $u = 1$, idle power at $u = 0$, and a linear $T(u)$ function. Various metrics based on $P(u)$ and $T(u)$ functions have been proposed to quantify and compare power consumption and/or energy efficiency of servers. We classify such server metrics broadly as 1) energy proportionality metric and 2) energy efficiency metric.

A. Energy Proportionality Metrics

Ever since Barrosso et al. [1] made a case for energy proportional computers, various metrics have been proposed to quantify energy proportionality. An ideal energy proportional server would consume power in proportion to its utilization, that is, $P(u)$ curve would be linear with $P(0) = 0$ (no power consumption when idle). Such a server will have constant energy efficiency throughout the utilization range. Energy proportionality has historically been measured using the Dynamic Range (DR) [14] metric,

$$\mathrm{DR} = \frac{P(1) - P(0)}{P(1)}. \tag{1}$$

A similar metric is Energy Proportionality Index (EPI),

$$\mathrm{EPI} = \frac{P(1) - P(0)}{P(1)} \cdot 100\%, \tag{2}$$

which is simply the dynamic range expressed as a percentage and was proposed by Mahadevan et al. [9], while a complementary metric, Idle-to-peak Power Ratio (IPR),

$$\mathrm{IPR} = \frac{P(0)}{P(1)}, \tag{3}$$

was proposed by Varsamopoulos et al. [18]. The problem with these metrics is that they only consider the end points in the $P(u)$ curve (power consumption at $u = 0$ and $u = 1$) and completely ignore the power consumption of a server at intermediate utilizations.

One of the most used metric [7], [8] to measure energy proportionality is the EP metric by Ryckbosch et al. [13].

$$\mathrm{EP} = 1 - \frac{\int_0^1 P(u)\,\mathrm{d}u - \int_0^1 P_{ideal}(u)\,\mathrm{d}u}{\int_0^1 P_{ideal}(u)\,\mathrm{d}u}, \tag{4}$$

where $P_{ideal}(u)$ is the power-utilization curve of an ideal energy proportional server with the same peak power. Since, $P_{ideal}(u)$ would be a line joining $P(0)$ to $P(1)$, we have $\int_0^1 P_{ideal}(u)\,\mathrm{d}u = \frac{P(1)}{2}$. Replacing this value in Eq. (4),

$$\mathrm{EP} = 2 - 2 \cdot \frac{\int_0^1 P(u)\,\mathrm{d}u}{P(1)} = 2 - 2 \cdot \int_0^1 P_N(u)\,\mathrm{d}u, \tag{5}$$

where $P_N(u)$ is the power-utilization curve normalized with respect to its peak power, $P(1)$. Therefore, EP is simply a scaled version of 1 minus the area under the $P_N(u)$ curve. Another recently proposed metric that takes into account the entire $P(u)$ curve is Energy Proportionality Coefficient (EPC) by Fiandrino et al. [2]. It is defined as,

$$\mathrm{EPC} = \int_0^1 sin2\alpha\,\mathrm{d}u, \tag{6}$$

where α is the angle made by the tangent of the $P(u)$ curve, $\alpha = tan^{-1}(\frac{\mathrm{d}P(u)}{\mathrm{d}u})$. However, since this metric just looks at the instantaneous slope, a $P(u)$ curve that increases in steps (a staircase function) will have $\mathrm{EPC} = 0$, an indication of no energy proportionality.

All of these metrics are compared and summarized in Table I. The following should be taken into consideration when using energy proportionality metrics for server comparison.

1) These metrics are normalized with respect to the peak power use. Thus, there is no way to know the actual energy use for a server with a given workload. Two servers with the same metric values could have very different actual energy consumptions.
2) These metrics do not consider throughput of a server. A server with a $P(u)$ curve that is flat ($\mathrm{EP} = 0$) could consume less power for the same throughput compared to a server with a $\mathrm{EP} = 1$ and a high peak power use.

B. Energy Efficiency Metrics

Energy efficiency of a data center is generally measured with Power Usage Effectiveness (PUE) [4], the ratio of total data center power consumption to the power consumption by IT equipment. Task schedulers are generally compared using the Energy-Delay Product (EDP) [10] metric. One particular benchmark called JouleSort [12], measures the energy required to sort a fixed number of records and uses SortedRecs/Joule as the metric. Another popular metric when comparing power-performance trade-off between computing systems is performance per Watt. The SPECpower benchmark [16] measures performance, in terms of ssj operations per second (ssj_ops), and power, in terms of Watts and the SPEC Overall Score (SOS) is calculated as the ratio of aggregate performance and aggregate power consumption at different utilization level,

$$\mathrm{SOS} = \frac{\sum_{i=0}^{10} T(i/10)}{\sum_{i=0}^{10} P(i/10)}. \tag{7}$$

Energy efficiency of servers are compared using this overall score. The continuous version for SOS could be calculated as,

$$\mathrm{SOS} = \frac{\int_0^1 T(u)\,\mathrm{d}u}{\int_0^1 P(u)\,\mathrm{d}u}. \tag{8}$$

TABLE I
DIFFERENT ENERGY PROPORTIONALITY METRICS

Metric	Formula	Range	Ideal value	higher/lower is better	Pros	Cons
Dynamic Range (DR) [14]	$\mathrm{DR} = \frac{P(1)-P(0)}{P(1)}$	[0,1]	1	higher	Easy to calculate	Only looks at end value in $P(u)$ curve
Energy Proportionality Index (EPI) [9]	$\mathrm{EPI} = \frac{P(1)-P(0)}{P(1)} \cdot 100\%$	[0,100]	100	higher	Easy to calculate	Only looks at end value in $P(u)$ curve
Idle-to-peak Power Ratio (IPR) [18]	$\mathrm{IPR} = \frac{P(0)}{P(1)}$	[0,1]	0	lower	Easy to calculate	Only looks at end value in $P(u)$ curve
Energy Proportionality (EP) [13]	$\mathrm{EP} = 2 - 2 \cdot \int_0^1 P_N(u)\,\mathrm{d}u$	[0,2]	1	higher	Takes entire P(u) curve into account	Does not consider the shape of $P(u)$ curve
Energy Proportionality Coefficient (EPC) [2]	$\mathrm{EPC} = \int_0^1 sin2\alpha\,\mathrm{d}u$	[0,1]	1	higher	Takes entire P(u) curve into account	Magnitude of $P(u)$ curve is ignored

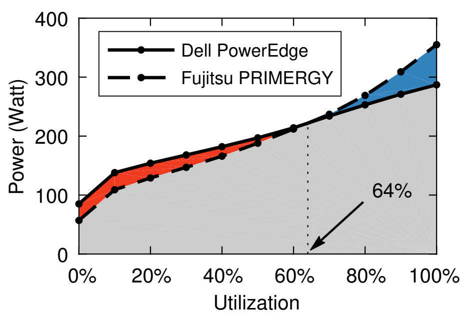

Fig. 2. Power versus utilization graph for two servers, Dell PowerEdge R7425 and Fujitsu Server PRIMERGY RX2540 M4, from SPECpower benchmark [16]. The two lines intersect at 64% utilization.

In practice, since $T(u)$ is generally a line joining 0 and $T(1)$ (peak throughput), we have $\int_0^1 T(u)\,\mathrm{d}u = \frac{T(1)}{2}$. Considering this assumption, we can establish a relation between SOS and EP from Eq. (5) and Eq. (8) as,

$$\mathrm{SOS} = \left(\frac{T(1)}{P(1)}\right)\left(\frac{1}{2-\mathrm{EP}}\right). \tag{9}$$

The above analytical relationship is an improvement on the approximate empirical relationship derived in [7]. Few observations can be made from this relation: 1) while the energy proportionality metric ignores peak power and throughput of servers, SOS takes them into account, and 2) given that the peak power and peak throughput for a server is fixed, SOS increases (and approaches infinity) as EP increases (and approaches 2).

III. Weighting Metrics with Workload Profile

In this section, we describe the importance of area under the power-utilization curve of a server and the notion of workload weighting to compare energy consumption between servers.

A. Area under the power-utilization curve

The energy usage of a server over time depends upon its power-utilization curve. In Figure 2, the power-utilization curve for two recent 2018 servers, Dell PowerEdge R7425 and Fujitsu Server PRIMERGY RX2540 M4, from the SPECpower benchmark [16] result, are shown. The two $P(u)$ curves intersect at 64% utilization. Assuming that the two systems have similar performance (throughput) over the utilization range, for server utilization of 64% and above, the Dell server consumes less power than the Fujitsu server. This implies that if the server utilization is always above 64%, then the Dell server consumes less energy than the Fujitsu server for the same workload. The opposite is true for server utilization below 64%. Hence, in terms of energy consumption, the Dell server is better than the Fujitsu server for utilization above 64% while being worse for utilization below 64%. So, which system is better overall? To answer this question, the total area under the two curves must be compared. In this particular case, the Fujitsu server (area of 197.2) is only slightly better than the Dell server (area of 199.7). However, an assumption that it is equally likely for the server to be at any of the utilization level is implied (if this were not the case, for example, if the servers were always utilized between 0% to 50%, then the Fujitsu server would clearly be better). Since, SOS and EP consider the notion of area under the $P(u)$ curve, these two metrics are taken into account for further analysis.

B. Workload weighting

In this paper, workload is defined as the rate at which a computing operation (for example, web request, transaction request, and database query) is offered to a server. The performance of a server is defined in terms of peak throughput, or the maximum rate at which a server can complete work, measured in operations per second (ops). It is important to note that same workload can create different levels of utilization of servers with different performance. For example, if a server is fully utilized at 1000 ops, another server with double the performance (twice as fast) will only be 50% utilized for the same workload. We assume this linear behavior, as in practice [16], throughput increases linearly with utilization of a server.

A time varying workload will result in a time varying utilization (and in turn, power use) of a server. Thus, energy consumption of a server will also depend on the workload. More specifically, it will depend on the probability density function (pdf) of the server utilization, $f(u)$. Since different organization/data centers will have different workload charac-

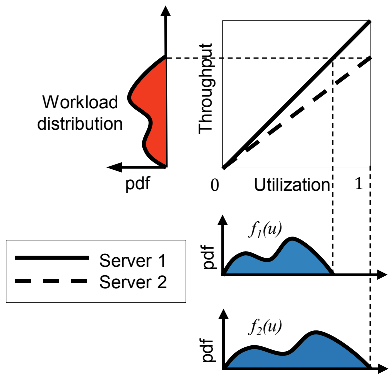

Fig. 3. Same workload distribution produces different utilization distribution for server 1 and server 2 with unequal performance.

teristics, an IT manager may want to select a server that is the most energy efficient for their particular workload. The EP or SOS metrics introduced in section II do not take workload into account. In the following section, we demonstrate that they do not correctly reflect the actual energy consumption of a server for different workloads. In order to make such a comparison, we propose workload weighted version of these metrics, WEP (Weighted EP) and WSOS (Weighted SOS),

$$\mathrm{WEP} = 2 - 2 \cdot \frac{\int_0^1 f(u) \cdot P(u)\, \mathrm{d}u}{P(1)}, \tag{10}$$

$$\mathrm{WSOS} = \frac{\int_0^1 f(u) \cdot T(u)\, \mathrm{d}u}{\int_0^1 f(u) \cdot P(u)\, \mathrm{d}u}. \tag{11}$$

Eq. (10) and Eq. (11) are the weighted counterpart of Eq. (5) and Eq. (8), respectively. In the special case when $f(u)$ is a uniform distribution, WEP and WSOS reduce to EP and SOS, respectively.

The workload weighted area under the $P(u)$ curve reflects the average power consumption by a server for that particular workload. We call this new metric the Weighted Average Power (WAP) of a server,

$$\mathrm{WAP} = \int_0^1 f(u) \cdot P(u)\, \mathrm{d}u. \tag{12}$$

The energy consumption by a server for a given workload is the product of WAP and time under consideration. Hence, WAP is a linear metric, that is, a server with 10% lower WAP than another server will consume 10% less energy compared to that server, for the same workload.

Two servers with different performance may have different $f(u)$ for the same workload distribution. Figure 3 shows an example of two servers where server 1 has a higher peak throughput than server 2. The pdf in left (in red) is the distribution of the workload offered to the servers. The same workload distribution is mapped to two different pdf at bottom (in blue), the server utilization distributions. More specifically, the utilization distribution of server 1, $f_1(u)$, is a shrunk version of the utilization distribution of server 2, $f_2(u)$. This illustrates that $f(u)$ captures both the workload characteristics as well as the throughput characteristics of a server assuming that the servers are capable of handling peak workload. Instead of single server comparison, if we need to find out which server (out of a heterogeneous pool of available servers) is most energy efficient for a data center (or a cluster), we propose to use $f_A(u)$, the distribution of average utilization over servers, assuming the data center (or the cluster) is composed of that same (homogeneous) server.

TABLE II
VARIANCE IN CHARACTERISTICS OF 42 SELECTED SPEC SERVERS

	Peak power (W)	Peak throughput (ssj_ops)	SOS (ssj_ops/W)	EP	Year
Min	241	3.16 M	8,274	0.6	2014
Median	288	3.55 M	11,540	0.9	2015
Max	469	5.36 M	13,398	1.0	2018

IV. Evaluation of Weighted Metrics

In this section, we evaluate and compare the usefulness of five different metrics in predicting the energy consumption of servers, 1) EP, 2) SOS, 3) WEP, 4) WSOS, and 5) WAP. We demonstrate the importance of considering the workload when selecting a server.

A. SPECpower performance-power data

SPEC maintains a database of results from SPECpower benchmark [16] reported by various hardware vendors like Dell, HP, Fujitsu, IBM, etc. SPEC reviews and regularly publishes results, and there are currently results for 560 servers out of which 40 are non-compliant with the standard. Since this data contains a wide range of servers with high variability in their characteristic (for example, we have servers with peak power consumption from 44.7 Watts to 6854 Watts), we filter out the data based on power and performance for further analysis. Specifically, out of the 520 valid results, we pick servers with peak power consumption below 500 Watts since most (61%) of the servers above this peak power were from multi node tests (average peak power of servers in real deployment is around 330 watts [14]). We pick servers with peak throughput above 3 million ssj operations per seconds (ssj_ops) as majority (93%) of the servers below this peak throughput are more than 5 years old (average lifetime of servers [14]). Such a selection makes sense as a data center operator would want to select from servers that are able to meet certain performance requirement while satisfying a power budget. We end up with 42 servers, and even though this is only 8% of the entire data, the variability in characteristics is significant as shown in Table II.

B. Google cluster utilization data

Google has publicly released a 29 day long cluster usage trace [21] of one of its cells consisting of about 12.5 thousand servers. The cluster usage trace has information on job and task events, their constraints and resource usage, attributes

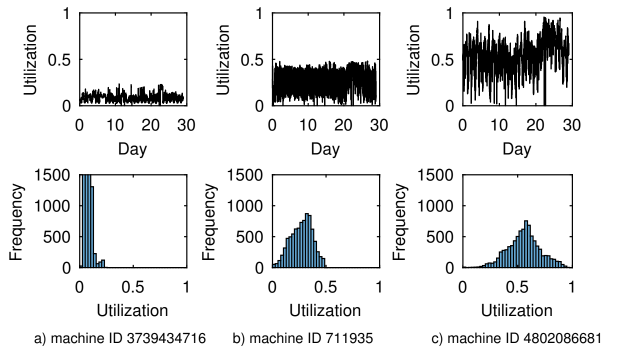

Fig. 4. CPU usage of some selected servers throughout the trace period of 29 days and their corresponding histogram. a) daily pattern is clearly visible where CPU usage does not exceed 0.25 b) no specific pattern where CPU usage does not exceed 0.5 c) weekly as well as daily patterns visible and CPU usage is high.

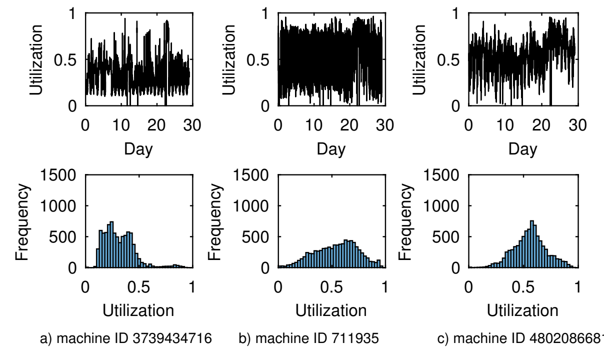

Fig. 5. CPU usage of same servers in Figure 4 and their corresponding histogram after scaling or removing from consideration. a) scaled by 4 b) scaled by 2 c) no scaling required.

of physical server, etc., but sensitive information has been obfuscated [11]. The data has various tables and our interest is in the "task_usage" table which logs information like physical server the task is running on, CPU usage, memory usage, etc. about each task in the cell for every measurement period of 5 minutes duration. Each of the 12,583 servers has a unique machine ID associated with it. The total CPU usage of a particular server for a 5 minute measurement interval is calculated by summing the CPU usage of all the tasks running on that particular server during the measurement interval. In this way, a time series of 8,351 data points for each server corresponding to average CPU usage in every 5 minute measurement interval over 29 days is constructed.

Figure 4 shows CPU utilization and the corresponding histogram of 3 selected servers. Daily or weekly patterns are visible in some of them. In the first figure, CPU utilization ranges from 0 to 0.25, in the second figure, it ranges from 0 to 0.5, while the last one ranges from 0 to 1. The "machine_events" table in the cluster trace provides information on server's normalized CPU and normalized memory capacity. We use this information to scale back CPU utilization such that it ranges from 0 to 1. For example, the server with machine ID 3739434716 (first server in Figure 4) has normalized CPU capacity of 0.25 and hence we observe its CPU usage is between 0 and 0.25. We scale its CPU usage by multiplying it by 4. There were negligibly few points (less than 0.001%) that were above 1 after scaling the CPU usage, which we truncated to 1. The overall average utilization of servers after scaling was about 40%, typical of such hyper-scale data center deployment [14], validating our approach. The "machine_events" table also contains information on when a server is added, removed, or updated in the cluster. We take into consideration only those servers that are present from the beginning of the trace and that are not updated or removed from the cluster throughout the 29 day trace period. There were 7,171 such servers (out of 12,583) that were in the cluster throughout the trace period.

Figure 5 shows CPU usage and the corresponding histogram of the same 3 selected servers in Figure 4 after being scaled. In our analysis, we multiply the CPU utilization trace of each Google server thus obtained by 3 million ssj_ops to get our workload trace for evaluation. This is because our set of 42 servers from SPEC data have peak throughput above 3 million ssj_ops. This avoids any overload situation (offered workload being greater than server's peak throughput).

C. Server selection with weighted metrics

Given one workload trace (from the Google cluster data) and one server characteristic (from SPECpower data), we can calculate the utilization and in turn the power consumed by that server. If we denote w to be the workload (measured in ssj_ops), server utilization can be calculated from the inverse of the throughput-utilization function of the server, $u = T^{-1}(w)$. Now, power can be calculated from this utilization using the power-utilization function, $P(u)$. Since, SPECpower data only has throughput and power at 11 discrete utilization levels, we linearly interpolate it to get $T(u)$ and $P(u)$ at all $0 \leq u \leq 1$. Energy consumption is the product of power and the time interval. Energy consumed by a server for the entire length of the workload trace can be calculated by summing up the energy consumed at each time interval.

We have a set of 42 different servers from the SPEC database, and each server may have a different energy consumption for a particular workload. This creates a ranking of servers from best to worst. The server that consumes the least energy for a workload is the best (most energy efficient) server for that workload. We have five metrics, two metrics that do not take workload into consideration (EP and SOS) and three workload weighted metrics (WEP, WSOS, and WAP). Best metric is the one that ranks the servers in a similar way as ranked according to their energy consumption (ground truth), that is, the metric is able to select the most energy efficient server, from a set of possible servers, given a workload. We need certain criteria to quantify which metric is better, and correlation has been found to be a good measure for this task [20]. If there is a high correlation between energy consumption of a server and the metric value, we can say that the metric is useful in determining energy efficiency. On the contrary, if the correlation is closer to zero, this implies that the metric does not inform on energy consumption of a server. We use three popular correlation coefficients for our evaluation, as common in the literature [20].

- Pearson's Linear Correlation Coefficient (PLCC):
$$\mathrm{PLCC} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{13}$$
where n is the number of observations of two variable $X = \{x_i\}_{1 \leq i \leq n}$ and $Y = \{y_i\}_{1 \leq i \leq n}$, $\bar{x}$ and $\bar{y}$ are sample mean of variables X and Y respectively.
- Spearman's Rank Correlation Coefficient (SRCC):
$$\mathrm{SRCC} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{14}$$
where d_i is difference between two ranks.
- Kendall's Rank Correlation Coefficient (KRCC):
$$\mathrm{KRCC} = \frac{2(n_c - n_d)}{n(n-1)} \tag{15}$$
where n_c and n_d are the number of concordant and discordant pairs, respectively.

While PLCC is a measure of linear dependency, SRCC and KRCC are a measure of rank correlation. Whenever we have some kind of subjective ranking and we need to measure how close this ranking is to the objective ranking (ground truth), we use SRCC and KRCC. However, for the sake of completeness, we have included PLCC in our analysis. Since we care more about ranking rather than the linear relation, SRCC and KRCC are more relevant for this study. SRCC and KRCC are both measuring the monotonicity of the two variables. We calculate PLCC, SRCC, and KRCC (higher magnitude is better) between the energy consumption of 42 servers (ground truth) with each of the five metric values for those same servers. Our objective is to find out which server would be the best for the given workload and which metric correctly predicts this. We take each of the 42 different SPEC servers, one at a time, and assume a data center is populated with that particular server (7,171 homogeneous servers in our case). Energy consumed by the data center given that a server was deployed can now be calculated separately for each of the 42 available server. In order to calculate the three workload weighted metrics, WEP, WSOS, and WAP, we consider the pdf of average utilization, $f_A(u)$ as described in the previous section. Note that $f_A(u)$ may be different for each of the 42 servers.

Figure 6 shows the metric versus energy consumption for all the five metrics. All three correlation coefficients for each of the metric is shown in Table III and the best value for each column is highlighted in bold. We find that although EP shows some negative PLCC, it is not a good metric to find the best server in terms of energy efficiency as SRCC and KRCC are both close to zero. Weighting this metric to get WEP does not affect the nature of the result as seen with small correlation values. SOS, which is an energy efficiency metric, seems to do better than EP in ranking of servers, but the correlation values are still small. However, when we weight this metric to get WSOS, results improve dramatically. This is seen both in Figure 6 (d) where WSOS almost monotonically increases as energy consumption decreases, and in Table III where all the correlation coefficients are close to -1. Similar results were found for WAP. Energy consumption increases as WAP increases, so there is a positive correlation. For comparison purposes such that all correlation coefficients are negative, the WAP metric values were multiplied with -1. This does not affect the observations or nature of results, but simply changes the sign of the correlation coefficient (we are interested in magnitude rather than sign). We can see that the WAP correlation coefficient values are also close to -1. Even though, the SRCC and KRCC values of WSOS are slightly higher than WAP, they are not statistically significantly different (using Fisher's r-to-z transformation and z-test at 0.05 level of significance). WSOS and WAP are able to correctly rank the servers according to energy consumption. In addition, WAP has a linear relation with energy consumption which is a desirable property.

TABLE III
CORRELATION OF SERVER ENERGY CONSUMPTION WITH DIFFERENT WEIGHTED AND UNWEIGHTED METRICS

Metric	PLCC	SRCC	KRCC
EP	-0.3581	0.0002	0.0221
SOS	-0.2033	-0.0563	-0.1220
WEP	-0.0259	0.0283	0.0197
WSOS	-0.9854	**-0.9880**	**-0.9233**
WAP	**-0.9937**	-0.9870	-0.9210

Next, we wish to determine the robustness of the metrics for different workloads. For this we consider the 7,171 workload traces individually and calculate PLCC, SRCC, and KRCC between the energy consumption of 42 servers with each of the five metric values given a particular workload trace. Since, we have 7,171 different workload traces, we repeat this calculation for each workload. Table IV lists the summary of correlation values (highest value in bold) of the five metrics for all the workloads. The Mann–Whitney U test between each 10 pairs of the five metrics repeated for all three correlation coefficient (30 tests in total), shows that all of them are statistically significantly different at 0.05 level of significance. Various observations can be made from this table:

- EP isn't a good metric to predict energy consumption of a sever for a workload as they have low correlation values.
- Weighting EP by the workload distribution to get WEP does not improve the result.
- SOS is better than EP at predicting energy consumption of servers but the correlation values are still small.
- WSOS and WAP are significantly better than EP, WEP, and SOS at predicting energy consumption of servers.
- WSOS has a higher range compared to WAP, signifying that WAP is more robust than WSOS and is also significantly better.

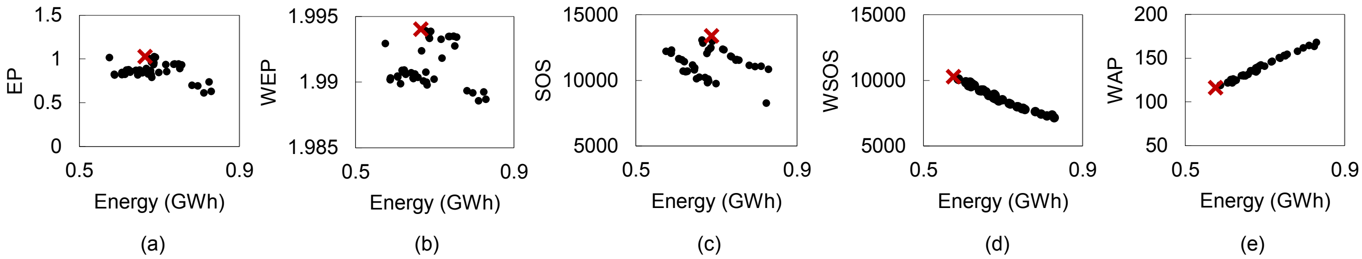

Fig. 6. Different metrics, a) EP, b) WEP, c) SOS, d) WSOS, and e) WAP, versus energy consumption for all 42 servers. The red cross shows the best server select by that particular metric.

TABLE IV
CORRELATION OF SERVER ENERGY CONSUMPTION WITH DIFFERENT WEIGHTED AND UNWEIGHTED METRICS FOR 7,171 WORKLOAD TRACES

Metric	PLCC				SRCC				KRCC			
	Min	Median	Mean	Max	Min	Median	Mean	Max	Min	Median	Mean	Max
EP	-0.6499	-0.3390	-0.3516	-0.2825	-0.2545	0.0184	-0.0080	0.0573	-0.1986	0.0267	0.0077	0.0499
SOS	-0.8318	-0.1616	-0.2016	0.0104	-0.8102	-0.0296	-0.0634	0.1017	-0.6446	-0.1034	-0.1223	-0.0012
WEP	-0.6981	-0.0104	-0.0385	0.0670	-0.5295	0.0385	0.0015	0.1010	-0.3868	0.0197	-0.0068	0.0592
WSOS	-0.9951	-0.9940	-0.9929	-0.5660	**-1.0000**	**-1.0000**	-0.9995	-0.7532	**-1.0000**	**-1.0000**	-0.9985	-0.6144
WAP	**-1.0000**	**-1.0000**	**-1.0000**	**-0.9980**	**-1.0000**	**-1.0000**	**-0.9999**	**-0.9906**	**-1.0000**	**-1.0000**	**-0.9993**	**-0.9419**

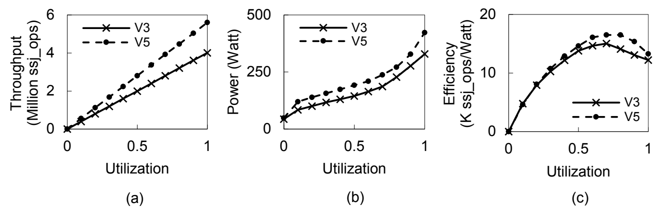

Fig. 7. Comparison between two servers that are considered best by SOS (Huawei Fusion Server 2288H V5) and WSOS/WAP (Huawei Fusion Server RH2288H V3). a) Throughput-utilization curve b) power-utilization curve c) efficiency-utilization curve.

TABLE V
COMPARISON OF TWO SERVERS, ONE RANKED HIGHEST BY SOS AND ANOTHER RANKED HIGHEST BY WEIGHTED METRIC

Server	Peak throughput	Peak power	SOS	WSOS	WAP	Energy (MWh)
V3	4,009,210	329	12,212	10,292	116	575.30
V5	5,604,530	423	13,398	8,486	141	687.34

D. *Example of selecting the wrong server*

We want to highlight why choosing a server ranked higher by the SOS metric without taking workload into consideration can result in choosing a server that would consume more energy than a server ranked lower. Among the set of 42 servers, the Huawei Fusion Server 2288H V5 has the highest SOS rating of 13,398 ssj_ops/W (shown with a red cross in Figure 6 (c)). While the server that consumed the least amount of energy was the Huawei Fusion Server RH2288H V3 (taking the entire 7,171 workloads for 29 days into consideration), which was marked the best by the weighted metrics WSOS and WAP (shown with a red cross in Figure 6 (d) and (e)).

Let us compare the two servers side-by-side to get a better understanding. We will refer to the two servers as V3 and V5 for brevity. Figure 7 shows the throughput, power, and efficiency curves for both the servers. We can observe that V5 has a higher throughput as well as power consumptions across the utilization range. V5 also has higher energy efficiency than V3 for most of the utilization range. Some important properties of both servers, along with metric values, are shown in Table V. V5 consumes 687.34 MWh energy, which is 20% more that V3's 575.30 MWh consumption, as shown in Figure 8, for the workload we considered. Even though V5 looks better in terms of power, performance, and efficiency, than V3, this counter intuitive result is due to the fact that the workload causes V5 to operate at lower utilization (and thus lower efficiency) than V3. More specifically, for the same workload, V5's overall average utilization was 20% while V3's overall average utilization was about 30%. And since, V3's energy efficiency at 30% utilization is higher than V5's energy efficiency at 20%, it consumes less energy overall. Our workload weighted metric was able to identify this relation as shown in Table V. When considering the 7,171 workload traces individually, V5 wasn't the best for any of them, consuming 5% to 27% (20% on average) higher energy than the best server. Selecting a more powerful and energy efficient server, by SOS metric, can cost 20% more electricity which highlights the importance of taking workload into consideration by data center operators.

V. RELATED WORK

Sound energy efficiency metrics are required to make comparisons and measure progress over time. Numerous metrics to evaluate data center efficiency including Energy Reuse Effectiveness (ERE) [17], Energy Reuse Factor (ERF) [17], Return Temperature Index (RTI) [6], and Rack Cooling Index (RCI) [6], have been proposed but the most used is still, PUE [4]. References [19] and [2] describe these, and many

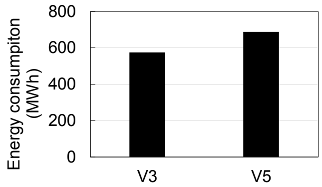

Fig. 8. Energy consumption of the cluster when using V3 and V5 servers.

other, metrics in detail. PUE, however, does not take energy efficiency of IT equipment into account, but only its power consumption. Metrics for system level energy efficiency like those proposed in [2], [9], [12], [13], [16], [18] are particularly of interest as reviewed in section II. Work in [8] studies the improvement of EP and SOS of server over the past decade using the SPECpower benchmark data. Their finding that high energy proportionality of a server doesn't necessarily imply higher energy efficiency is in concert with our findings. Similarly, [7] studies different energy proportionality metrics and tries to find an empirical, if not an analytical, relation between them. They find an empirical relation between EP and SOS, which is approximate. We believe our relationship in Eq. (9) is exact, assuming servers have a linear throughput-utilization curve. Hanson et al. [5] emphasized the importance of capturing the workload variability for power-performance benchmarks, but did not study or quantify workload weighting.

Our work differs from, and builds upon, existing work in two ways. Our work reveals new insights into the effect of workload on power-performance metrics. Our work also develops and empirically evaluates how weighting existing and new metrics with workload can result in a better selection of the most energy efficient server for given workload.

VI. Summary and Future Work

In summary, this paper puts forward the idea of workload weighting of existing energy efficiency metrics for predicting the most energy efficient server. New workload weighted metrics are proposed and empirically evaluated to show their reliability in ranking of servers for various workloads. This paper also demonstrates how ignoring workload characteristics can lead to selection of a server that consumes 20% more energy than the most energy efficient server for a particular Google cluster workload. While there are numerous features of a server to take into consideration, we believe that energy consumption along with performance are the two key features and our workload weighted metrics will help data center operators in selecting the appropriate server for their data center.

Publicly available real data center server utilization data is very limited. The publicly available Google cluster usage data was transformed to obtain individual server utilizations used as workload in our work. Validating results in this paper with representative server utilization data from smaller enterprise and/or service provider type data centers is one possible future direction. In the future, we would like to study the effect of different types of workload (CPU intensive, memory intensive, or network intensive) on our results.

Acknowledgement

We would like to thank the anonymous reviewers for their valuable feedback that helped improve the quality of the paper.

References

[1] L. A. Barroso and U. Hölzle, "The case for energy-proportional computing," *Computer*, vol. 40, pp. 33–37, 2007.
[2] C. Fiandrino, D. Kliazovich, P. Bouvry, and A. Y. Zomaya, "Performance and energy efficiency metrics for communication systems of cloud computing data centers," *IEEE Transactions on Cloud Computing*, vol. 5, no. 4, pp. 738–750, 2017.
[3] H. Fuchs, A. Shehabi, M. Ganeshalingam, L.-B. Desroches, B. Lim, K. Roth, and A. Tsao, "Characteristics and energy use of volume servers in the united states," Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 2017.
[4] Green Grid, "Green grid data center power efficiency metrics: PUE and DCiE," Tech. Rep., 2007.
[5] H. Hanson, K. Rajamani, J. Rubio, S. Ghiasi, and F. Rawson, "Benchmarking for power and performance," in *Proceedings of SPEC Benchmarking Workshop*, 2007.
[6] M. K. Herrlin, "Airflow and cooling performance of data centers: Two performance metrics," *ASHRAE transactions*, vol. 114, no. 2, pp. 182–187, 2008.
[7] C.-H. Hsu and S. W. Poole, "Measuring server energy proportionality," in *Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering*, 2015, pp. 235–240.
[8] C. Jiang, Y. Wang, D. Ou, B. Luo, and W. Shi, "Energy proportional servers: Where are we in 2016?" in *2017 IEEE 37th International Conference on Distributed Computing Systems*, 2017, pp. 1649–1660.
[9] P. Mahadevan, P. Sharma, S. Banerjee, and P. Ranganathan, "A power benchmarking framework for network devices," in *NETWORKING 2009*. Springer Berlin Heidelberg, 2009, pp. 795–808.
[10] A.-C. Orgerie, M. D. d. Assuncao, and L. Lefevre, "A survey on techniques for improving the energy efficiency of large-scale distributed systems," *ACM Comput. Surv.*, vol. 46, no. 4, pp. 47:1–47:31, 2014.
[11] C. Reiss, J. Wilkes, and J. L. Hellerstein, "Google cluster-usage traces: format + schema," Google Inc., Mountain View, CA, USA, Tech. Rep., 2011, revised 2014-11-17 for version 2.1.
[12] S. Rivoire, M. A. Shah, P. Ranganathan, and C. Kozyrakis, "Joulesort: A balanced energy-efficiency benchmark," in *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '07, 2007, pp. 365–376.
[13] F. Ryckbosch, S. Polfliet, and L. Eeckhout, "Trends in server energy proportionality," *Computer*, vol. 44, no. 9, pp. 69–72, 2011.
[14] A. Shehabi, S. Smith, D. Sartor, R. Brown, M. Herrlin, J. Koomey, E. Masanet, N. Horner, I. Azevedo, and W. Lintner, "United states data center energy usage report," Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 2016.
[15] SPEC Power Committee, "SPEC power and performance benchmark methodology," Standard Performance Evaluation Corporation, Tech. Rep., 2014, version 2.2.
[16] Standard Performance Evaluation Corporation. (2008) SPECpower_ssj 2008. [Online]. Available: https://www.spec.org/power_ssj2008/
[17] B. Tschudi, O. Vangeet, J. Cooley, and D. Azevedo, "ERE: A metric for measuring the benefit of reuse energy from a data center," *White Paper*, 2010.
[18] G. Varsamopoulos and S. K. Gupta, "Energy proportionality and the future: Metrics and directions," in *2010 39th International Conference on Parallel Processing Workshops*, 2010, pp. 461–467.
[19] L. Wang and S. U. Khan, "Review of performance metrics for green data centers: a taxonomy study," *The journal of supercomputing*, vol. 63, no. 3, pp. 639–656, 2013.
[20] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.
[21] J. Wilkes, "More Google cluster data," Google research blog, Nov. 2011, posted at http://googleresearch.blogspot.com/2011/11/more-google-cluster-data.html.